

A new method of kernel-smoothing estimation of the ROC curve

Michał Pulit¹

Received: 13 June 2013 / Published online: 28 December 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract The receiver operating characteristic (ROC) curve is a popular graphical tool for describing the accuracy of a diagnostic test. Based on the idea of estimating the ROC curve as a distribution function, we propose a new kernel smoothing estimator of the ROC curve which is invariant under nondecreasing data transformations. We prove that the estimator has better asymptotic mean squared error properties than some other estimators involving kernel smoothing and we present an easy method of bandwidth selection. By simulation studies, we show that for the limited sample sizes, our proposed estimator is competitive with some other nonparametric estimators of the ROC curve. We also give an example of applying the estimator to a real data set.

Keywords ROC curve · Nonparametric estimation · Kernel smoothing · Bandwidth selection

Mathematics Subject Classification 62G05 · 62G20

1 Introduction

The receiver operating characteristic (ROC) curve is used to describe the performance of a diagnostic test, which on the basis of some observable measurements, assigns individuals to one of two different groups. For definiteness, let us think of them as the groups of diseased and healthy patients. This medical terminology is related to the fact that, in practice, the ROC curves are mainly used in medicine. However, their applications were recently extended to many other fields like economics and data

✉ Michał Pulit
michal.pulit@pwr.edu.pl

¹ Wrocław University of Technology, ul. Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

mining. More information about the ROC curves and their possible applications can be found, for example, in [Swets \(1988\)](#), [Pepe \(2003\)](#) and [Krzanowski and Hand \(2009\)](#). For a given cutoff point $c \in \mathbb{R}$, let an individual be classified as healthy if its test score is greater than c and as diseased otherwise. Suppose that the real random variables X and Y denote the test score in the groups of healthy and diseased individuals, respectively, and let $F(x) = P(X \leq x)$ and $G(x) = P(Y \leq x)$ be their continuous and strictly increasing distribution functions. The accuracy of the test is typically summarized by the sensitivity and specificity, given by $SE(c) = 1 - G(c)$ and $SP(c) = F(c)$, respectively. The ROC curve is a plot of $SE(c)$ versus $1 - SP(c)$ for all possible cutoff values $c \in \mathbb{R} \cup \{-\infty, \infty\}$. Equivalently, it can be defined as

$$R(t) = 1 - G(F^{-1}(1 - t)), \quad t \in [0, 1]. \quad (1)$$

Let $\mathbf{X}_m = (X_1, \dots, X_m)$ and $\mathbf{Y}_n = (Y_1, \dots, Y_n)$ be independent simple samples from healthy and diseased populations, respectively, and let F_m and G_n denote their empirical cumulative distribution functions. The most commonly used nonparametric estimator of $R(t)$ is the empirical ROC curve, which is of the form

$$R_{m,n}(t) = 1 - G_n(F_m^{-1}(1 - t)), \quad t \in [0, 1]. \quad (2)$$

Asymptotic properties of this estimator were studied by [Hsieh and Turnbull \(1996\)](#). Among other things they proved that, under some basic assumptions for F and G , $R_{m,n}(t)$ converges to the true ROC curve uniformly on $[0, 1]$ with probability one.

Although the empirical ROC curve is very simple and very popular, its obvious weakness is being a step function, while $R(t)$ is continuous and smooth. One of the ways to obtain a continuous estimator of $R(t)$ is to use the kernel smoothing method. [Zou et al. \(1997\)](#) proposed a nonparametric estimator of $R(t)$ from kernel estimates for the density functions of F and G . [Lloyd \(1998\)](#), using kernel estimates directly for F and G , obtained a smooth ROC curve estimator given by

$$\tilde{R}_{m,n}(t) = 1 - \tilde{G}_n(\tilde{F}_m^{-1}(1 - t)), \quad t \in [0, 1], \quad (3)$$

where

$$\tilde{F}_m(x) = \frac{1}{m} \sum_{j=1}^m \mathcal{Q}\left(\frac{x - X_j}{h_F}\right), \quad \tilde{G}_n(x) = \frac{1}{n} \sum_{j=1}^n \mathcal{Q}\left(\frac{x - Y_j}{h_G}\right)$$

are kernel estimators of F and G with a kernel function \mathcal{Q} , $\mathcal{Q}(v) = \int_{-\infty}^v \mathcal{Q}(z) dz$ and bandwidths h_F and h_G . [Lloyd and Zhou \(1999\)](#) proved that estimator (3) has better asymptotic mean squared error (MSE) properties than the empirical ROC curve. Unfortunately, to the best of our knowledge, in the case of estimator (3), there is no uniform, but only pointwise convergence to $R(t)$. Moreover, the kernel ROC curve estimator is not invariant under monotone data transformations, which may be undesirable in some practical applications. The problem of transformation-invariant nonparametric estimation of the ROC curve is considered, e.g., in [Du and Tang \(2009\)](#) and [Tang et al. \(2010\)](#). Finally, estimator (3) involves two separate bandwidth parameters, so

special care is required for bandwidth selection (Zhou and Harezlak 2002; Hall and Hyndmann 2003).

To overcome some of the mentioned drawbacks, different methods of smoothing the empirical ROC curve were proposed, including local linear smoothing (Peng and Zhou 2004), Bayesian bootstrap (Gu et al. 2008) and bandwidth-free smoothing of the empirical CDFs (Jokiel-Rokita and Pulit 2012). In this paper, instead of estimating the ROC curve as the composition of estimators of F^{-1} and G , we use the fact that for $Z = 1 - F(Y)$,

$$P(Z \leq t) = P(Y > F^{-1}(1 - t)) = 1 - G(F^{-1}(1 - t)) = R(t), \quad (4)$$

and propose to estimate $R(t)$ as the cumulative distribution function of Z . It is clear that without any knowledge about F , we need to obtain a predictor of the unknown random sample $\mathbf{Z}_n = (1 - F(Y_1), \dots, 1 - F(Y_n))$. The simplest way to do this is to substitute the unknown distribution function F by its any estimator \hat{F} . Based on the vector $\hat{\mathbf{Z}}_n = (1 - \hat{F}(Y_1), \dots, 1 - \hat{F}(Y_n))$, we can directly estimate $R(t)$, using the well known method of the kernel distribution function estimation.

In Sect. 2 we define a new kernel smoothing estimator of the ROC curve, which is invariant to nondecreasing data transformations and involves only one bandwidth parameter. We also show some asymptotic results, including a MSE comparison of the proposed estimator and the kernel-smoothed estimator proposed by Lloyd (1998). In Sect. 3 we propose a method of bandwidth selection. Section 4 contains results of simulation studies. Finally, in Sect. 5 we apply the proposed estimator to real data. All proofs are put in Appendices 1 and 2.

2 Main results

Let $\mathbf{X}_m = (X_1, \dots, X_m)$ and $\mathbf{Y}_n = (Y_1, \dots, Y_n)$ be independent simple samples from unknown distribution functions F and G with the same supports $I_F = I_G \subseteq \mathbb{R}$ and with density functions f and g (with respect to Lebesgue measure), respectively. Let K be a continuous symmetric density function with support $[-1, 1]$ and denote $\mathcal{K}(x) = \int_{-\infty}^x K(y)dy$. Define the smooth ROC curve estimator as

$$\hat{R}_{m,n}(t) = \frac{1}{n} \sum_{i=1}^n \mathcal{K} \left(\frac{t - 1 + F_m(Y_i)}{h_n} \right), \quad (5)$$

where $h_n > 0$ is a bandwidth parameter and F_m denotes the empirical distribution function of the sample \mathbf{X}_m . For the estimator $\hat{R}_{m,n}$ to have better (than some other estimators) asymptotic MSE properties, the kernel function K should satisfy some conditions like e.g. $\int_{-1}^1 K''(x)dx < 0$. Therefore, for simplicity, we assume that K is the Epanechnikov kernel $K(x) = 3/4(1 - x^2)\mathbf{1}_{[-1,1]}(x)$.

If we consider the ROC curve as the distribution function of the random variable $Z = 1 - F(Y)$ [see (4)], it is easily seen that $\hat{R}_{m,n}(t)$ is the kernel distribution function

estimator based on $\hat{Z}_i = 1 - F_m(Y_i)$, which are estimators of $Z_i = 1 - F(Y_i)$, $i = 1, 2, \dots, n$.

Remark 1 If we apply the same nondecreasing transformation to samples \mathbf{X}_m and \mathbf{Y}_n , the estimator given by (5) does not change, which means that $\hat{R}_{m,n}(t)$ is transformation invariant. Therefore, without loss of generality, we can assume that $I_F = I_G = \mathbb{R}$.

Theorem 1 Assume that $R''(s)$ exists for s near $t \in (0, 1)$, $R''(s)$ is continuous at $s = t$ and let $h_n \rightarrow 0$. Then

$$\text{Bias}(\hat{R}_{m,n}(t)) = \frac{R''(t)}{10} h_n^2 + O\left(\frac{1}{m} + \frac{1}{m^2 h_n^2}\right) + o(h_n^2), \quad (6)$$

$$\begin{aligned} \text{Var}(\hat{R}_{m,n}(t)) = & \frac{R(t)(1-R(t))}{n} + \frac{t(1-t)[R'(t)]^2}{m} - \frac{\frac{9}{35}R'(t)\left[\frac{m}{n} - R'(t)\right]h_n}{m} \\ & - \frac{3t^2(1-t)^2[R'(t)]^2}{m^2 h_n^2} + \frac{\frac{15}{4}t^3(1-t)^3[R'(t)]^2}{m^3 h_n^4} \\ & + O\left(\frac{m+n}{m^2 n h_n} + \frac{m+n}{m^3 n h_n^3} + \frac{m+n h_n}{m^4 n h_n^5}\right) + o\left(\frac{h_n}{m}\right). \end{aligned} \quad (7)$$

Let $k_n \in \mathbb{N}$ denote the minimal sample size for which the MSE of the empirical ROC curve is no greater than the MSE of estimator (3), based on a sample of size $n \in \mathbb{N}$. Lloyd and Zhou (1999) showed that, under some assumptions on the kernel function and the bandwidths h_F and h_G , the difference $k_n - n$ is divergent to infinity and $k_n - n \sim n\sqrt{h_F h_G}$. The proposed estimator $\hat{R}_{m,n}$ has an analogous advantage, not only over the empirical ROC curve, but also over estimator (3) proposed by Lloyd (1998). Assume that

$$m = m(n) = n\lambda_n, \quad \lambda_n \rightarrow \lambda \in (0, \infty), \quad (\text{M})$$

and denote

$$b_n(t) = \min \left\{ j \in \mathbb{N} : \text{MSE}(\tilde{R}_j^*(t)) \leq \text{MSE}(\hat{R}_n(t)) \right\}, \quad (8)$$

where \tilde{R}_n^* is the kernel ROC curve estimator given by (3), with the asymptotic optimal (in the sense of minimizing the MSE) bandwidths h_F^* and h_G^* , which are $O(m^{-1/3})$ and $O(n^{-1/3})$, respectively (Lloyd and Zhou 1999; Hall and Hyndmann 2003). The asymptotic bias and variance of $\tilde{R}_{m,n}$ are given by

$$\begin{aligned} \text{b}(\tilde{R}_n(t)) = & \frac{\alpha R''(t)}{2} \left[h_F^2 f^2(F^{-1}(1-t)) + \frac{t(1-t)}{m} \right] \\ & + \frac{1}{2}(h_F^2 - h_G^2)g'(F^{-1}(1-t)) + o\left(h_F^2 + \frac{1}{m}\right), \end{aligned} \quad (9)$$

$$\text{Var}(\tilde{R}_n(t)) = \frac{R(t)[1-R(t)]}{n} + \frac{t(1-t)[R'(t)]^2}{m} + O\left(\frac{h_G}{n}\right), \quad (10)$$

and were derived by [Lloyd \(1998\)](#). Therefore, using the condition [\(M\)](#), we get

$$\text{MSE} \left(\tilde{R}_n^*(t) \right) = \frac{R(t) [1 - R(t)]}{n} + \frac{t(1-t)[R'(t)]^2}{\lambda_n n} + O \left(n^{-\frac{4}{3}} \right). \quad (11)$$

Theorem 2 Suppose that assumptions of [Theorem 1](#) hold and the condition [\(M\)](#) is satisfied. Then, if $nh_n^2 \rightarrow \infty$ and $nh_n^4 \rightarrow 0$, we have

$$\lim_{n \rightarrow \infty} \frac{b_n(t)}{n} = 1.$$

Under the additional assumptions $nh_n^3 \rightarrow 0$ and $\lambda - \lambda_n = O(n^{-1/3})$, we get

$$\lim_{n \rightarrow \infty} h_n^2 (b_n(t) - n) = \frac{3t^2(1-t)^2[R'(t)]^2}{R(t)[1-R(t)]\lambda^2 + t(1-t)[R'(t)]^2\lambda} \geq 0,$$

and the above limit is strictly positive if $R'(t) > 0$. Finally, if we assume that $nh_n^2 \rightarrow \delta \in (0, \infty)$ and $\delta > \frac{5}{4\lambda}t(1-t)$, then

$$\lim_{n \rightarrow \infty} \frac{b_n(t)}{n} = \frac{\delta^2}{\delta^2 - \frac{3t^2(1-t)^2[R'(t)]^2 \left\{ \lambda\delta - \frac{5}{4}t(1-t) \right\}}{R(t)[1-R(t)]\lambda^3 + t(1-t)[R'(t)]^2\lambda^2}} \geq 1,$$

and the above limit is strictly positive if $R'(t) > 0$.

Remark 2 The MSE of the ROC curve estimator proposed by [Peng and Zhou \(2004\)](#), with asymptotically optimal choice of bandwidth, has the same form as the MSE of the estimator proposed by [Lloyd \(1998\)](#), given by [\(11\)](#) (see [Peng and Zhou \(2004\)](#), Sect. 3). Therefore, [Theorem 2](#) remains true if, instead of the estimator \tilde{R}_n^* appearing in [definition \(8\)](#) of $b_n(t)$, we insert the Peng and Zhou's estimator.

3 Bandwidth selection

In this section we deal with the issue of choosing the parameter h_n , appearing in [\(5\)](#). In the problem of bandwidth selection when estimating the distribution function, to the best of our knowledge, only two methods have been investigated: plug-in and cross-validation. The plug-in bandwidth choice was studied e.g. by [Altman and Leger \(1995\)](#) and [Polansky and Baker \(2000\)](#). The least-squares cross-validation method was analyzed in [Sarda \(1993\)](#) and in [Bowman et al. \(1998\)](#). It seems that an idea presented in the last paper may be adapted to our problem. Bowman et al. proposed the method which minimizes the function

$$CV_0(h) = \frac{1}{n} \sum_{i=1}^n \int (I(x - x_i) - \tilde{F}_{n,-i}(x, h))^2 dx, \quad (12)$$

where $I(x - x_i) = 1$ if $x - x_i \geq 0$ and 0 in other case, and $\tilde{F}_{n,-i}(x, h)$ denotes the kernel distribution function estimator constructed from the data with observation x_i omitted. Analogously, one can choose the bandwidth parameter h_n by minimizing the function

$$CV(h) = \frac{1}{n} \sum_{i=1}^n \int_0^1 \left(I(t - 1 + F_m(y_i)) - \hat{R}_{m,n,-i}(t) \right)^2 dt, \quad (13)$$

where

$$\hat{R}_{m,n,-i}(t) = \frac{1}{n-1} \sum_{j \neq i} \mathcal{K} \left(\frac{t - 1 + F_m(y_j)}{h_n} \right). \quad (14)$$

This method of bandwidth selection works decently and usually leads to the estimator $\hat{R}_{m,n}$ with the MSE smaller than in the case of the empirical ROC curve. However, when the sample sizes are small, it is not very stable and often gives too small or too large parameters h_n , which results in under- or oversmoothed estimated ROC curves, respectively. Moreover, procedure of numerical minimization of the function CV , repeated many times, is time consuming

For that reason we propose another method of choosing the parameter h_n . From Theorem 2 it follows that for fixed $t \in (0, 1)$ and for $nh_n^2 \rightarrow \delta$, where $\frac{5}{4\lambda}t(1-t) < \delta < \infty$, we have

$$\lim_{n \rightarrow \infty} \frac{b_n}{n} = \frac{\delta^2}{\delta^2 - \frac{3t^2(1-t)^2[R'(t)]^2 \left\{ \lambda\delta - \frac{5}{4}t(1-t) \right\}}{R(t)[1-R(t)]\lambda^3 + t(1-t)[R'(t)]^2\lambda^2}} := \Psi(\delta) \geq 1,$$

and it is easy to check that the function $\Psi(\delta)$ is maximized for

$$\delta^* = \frac{5t(1-t)}{2\lambda}.$$

Therefore, for fixed $t \in (0, 1)$, to maximize the asymptotic relative efficiency of $\hat{R}_n(t)$ with respect to $\tilde{R}_n^*(t)$, the bandwidth parameter h_n should be selected in such a way that $nh_n^2 \rightarrow \delta^*$. Hence, we propose to choose the bandwidth parameter which depends on t and is of the form

$$h_n^*(t) = c_n \sqrt{\frac{\delta^*}{n}} = \frac{c_n \sqrt{5t(1-t)}}{\sqrt{2n\lambda}}, \quad (15)$$

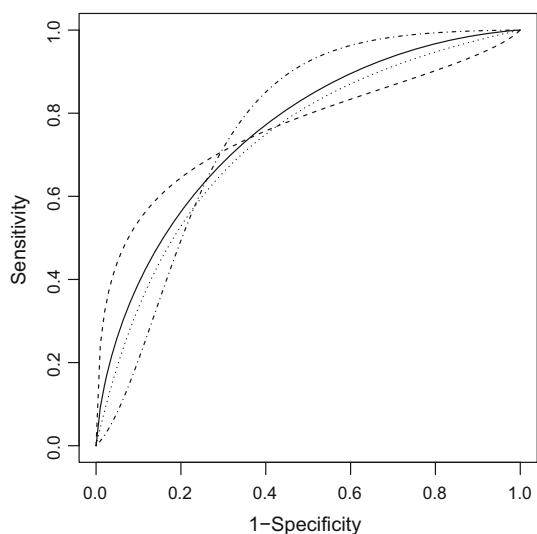
where c_n is some sequence converging to 1. Note that our proposed method of bandwidth selection gives the smoothing parameter $h_n^*(t)$ which, in contrast to the optimal bandwidth(s) obtained by other methods relating to some other kernel ROC curve estimators (e.g. [Lloyd and Zhou 1999](#); [Hall and Hyndmann 2003](#); [Peng and Zhou 2004](#)), does not depend on the unknown distribution functions F and G . Therefore the parameter $h_n^*(t)$ is easy to compute. Moreover, $h_n^*(t)$ becomes small near the ends of the interval $[0, 1]$, which results in a reduction of the bias of the proposed estimator, especially for t close to 0.

4 Simulation study

A small simulation study was performed to investigate the efficiency of the proposed estimator of the ROC curve for the limited sample sizes. We considered four different combinations of the distribution functions. In the first two studies, both F and G belong to the same family of distributions, normal or logistic. The parameters are selected so that the resulting ROC curves have similar shapes (see Fig. 1). In the other two studies, F and G are different: if one is normal, the other is logistic. In this case also the corresponding ROC curves are completely different. In the simulations we used 1000 samples of equal sizes $m = n = 20, 50$. For each of the considered ROC curves and sample sizes, we computed the empirical ROC curve (EM), the smoothed empirical ROC curve (SEM) of Joki-Rokita and Pulit (2012), based on smoothed empirical CDFs, the Bayesian bootstrap estimator (BB) proposed by Gu et al. (2008), the local linear smoothing estimator (LLS) of Peng and Zhou (2004), the Lloyd's kernel-smoothed estimator (KS) and the new kernel smoothing estimator (NKS) proposed in this paper.

Although the choice of the sequence c_n appearing in (15) does not affect on the asymptotic behavior of the estimator $\hat{R}_{m,n}$, for the limited sample sizes the best results are achieved when $c_n \approx 1.5 - 2.5$, depending on the estimated ROC curve and the value of n . In the simulation study, choosing $h_n^*(t)$ for our estimator, for simplicity, we decided to take $c_n = 1 + 1.8n^{-1/5}$ in all the considered cases. In the problem of bandwidth selection for the kernel estimator (KS), we used the normal-reference method proposed by Hall and Hyndmann (2003), which is recommended when the sampled distributions are not far from normal. The authors found that in the context of the ROC curve estimation, proposed method give substantial improvement in the mean integrated squared error over other known methods of bandwidth selection. Finally, in the case of the local linear smoothing estimator (LLS), we choose the smoothing

Fig. 1 The ROC curve corresponding to: (1) $X \sim \mathcal{N}(0, 1), Y \sim \mathcal{N}(1, 1)$ (solid curve); (2) $X \sim \mathcal{LG}(0, 2), Y \sim \mathcal{LG}(3, 2)$ (dotted curve); (3) $X \sim \mathcal{LG}(0, 1), Y \sim \mathcal{N}(2.5, 9)$ (dashed curve); (4) $X \sim \mathcal{N}(0, 9), Y \sim \mathcal{LG}(2.5, 1)$ (dot-dashed curve)



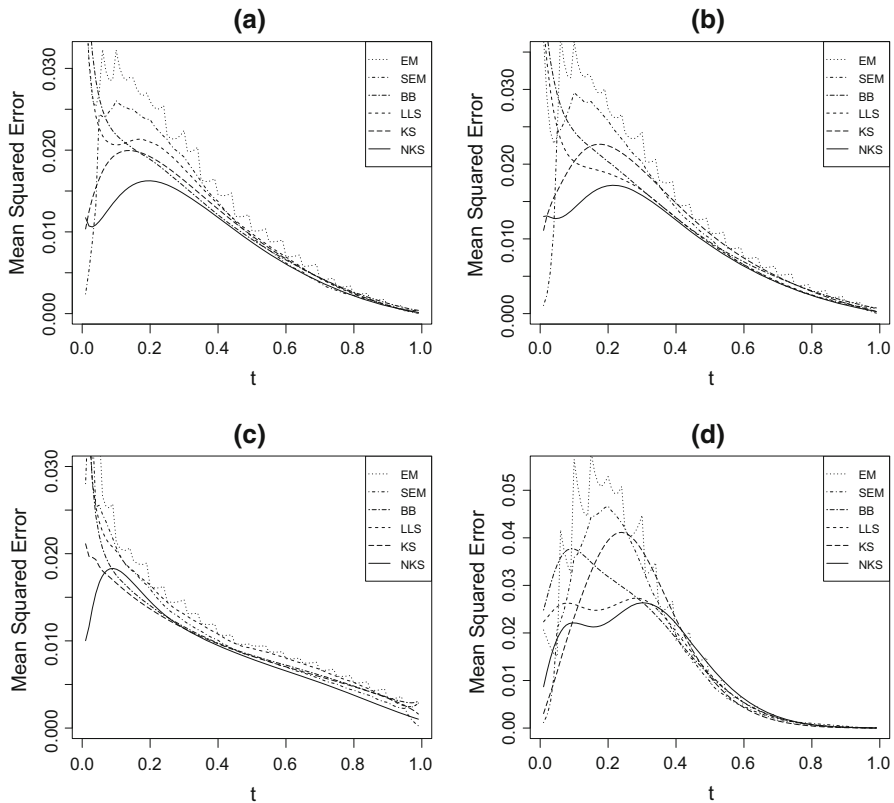


Fig. 2 The graphs of estimated MSE on the unit interval for the sample sizes $m = n = 20$ and for data from: **a** $F \sim \mathcal{N}(0, 1)$, $G \sim \mathcal{N}(1, 1)$; **b** $F \sim \mathcal{L}\mathcal{G}(0, 2)$, $G \sim \mathcal{L}\mathcal{G}(3, 2)$; **c** $X \sim \mathcal{L}\mathcal{G}(0, 1)$, $Y \sim \mathcal{N}(2.5, 9)$; **d** $X \sim \mathcal{N}(0, 9)$, $Y \sim \mathcal{L}\mathcal{G}(2.5, 1)$

parameter which minimizes the mean trimmed integrated squared error, assuming knowledge of the distribution functions F and G , [see Peng and Zhou (2004), Sect. 3].

Figures 2 and 3 display the results of the simulations for the sample sizes $m = n = 20$ and $m = n = 50$, respectively. Every figure contains four plots corresponding to four different ROC curves which are to be estimated (see Fig. 1) and every single plot compares the considered ROC curve estimators in term of their mean squared error (MSE) on the unit interval. The obtained results indicate that the proposed estimator (NKS) is competitive with other estimators, also for the limited sample sizes. In the problem of estimation of the ROC curve it performs better than the empirical ROC curve (EM), the smoothed empirical ROC curve (SEM) and the Bayesian bootstrap estimator (BB). In some of the cases it is also better than two other estimators, (KS) and (LLS).

Supplementary materials to the paper, containing some box-plots comparing the accuracy of the estimators in term of MSE when estimating AUC, are available at <https://drive.google.com/file/d/0B3L4pdDwuWxvT0RRbmUzWGtaa28/view?pli=1>.

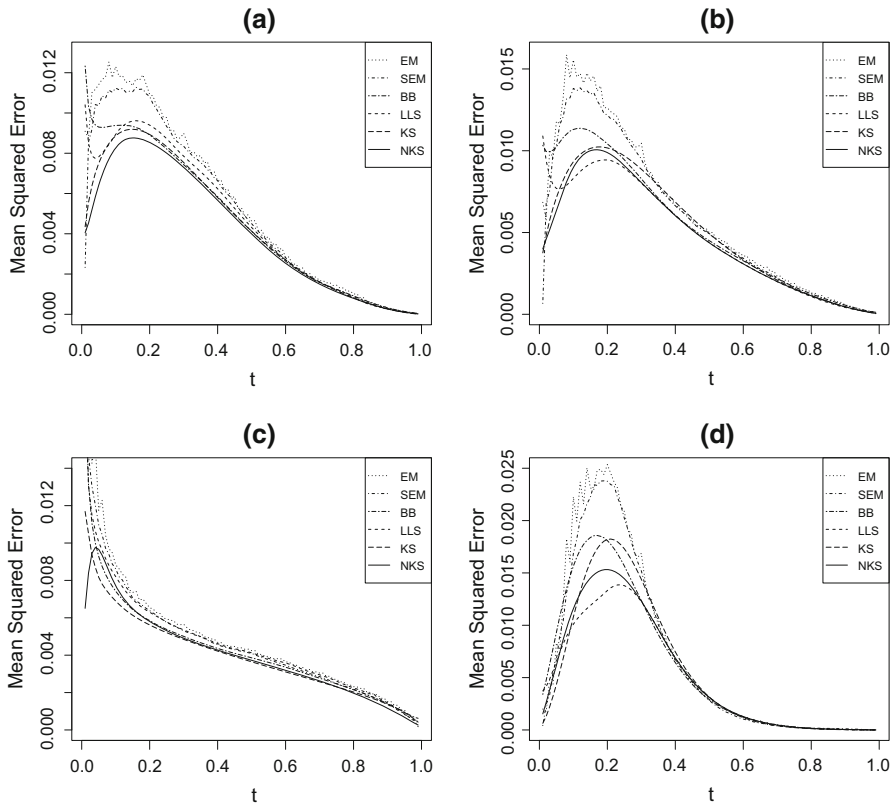


Fig. 3 The graphs of estimated MSE on the unit interval for the sample sizes $m = n = 50$ and for data from: **a** $F \sim \mathcal{N}(0, 1)$, $G \sim \mathcal{N}(1, 1)$; **b** $F \sim \mathcal{LG}(0, 2)$, $G \sim \mathcal{LG}(3, 2)$; **c** $X \sim \mathcal{LG}(0, 1)$, $Y \sim \mathcal{N}(2.5, 9)$; **d** $X \sim \mathcal{N}(0, 9)$, $Y \sim \mathcal{LG}(2.5, 1)$

5 Real data analysis

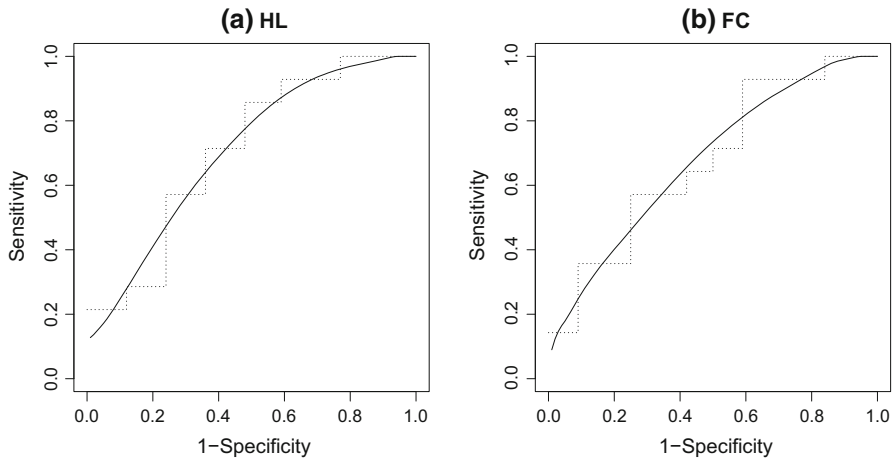
To illustrate our method, we apply it to the set of real data which comes from a clinical study performed from November 2008 to August 2011 by a research team led by Dr. Krzysztof Tupikowski from Department of Urology and Oncological Urology, Wrocław Medical University.

One investigated the efficacy of combined treatment of interferon alpha and metronomic cyclophosphamide in patients with metastatic kidney cancer not eligible for tyrosine kinase inhibitors treatment with various negative prognostic factors for survival. It has been approved by an independent local bioethics committee. One of the secondary goals of the study was to assess if there are any predictive factors for response to this novel combination treatment.

Table 1 contains presence (1) (or absence - 0) of clinical response (CR) observed at 24-th week of treatment, hemoglobin level (HL) and serum fibrinogen concentration (FC) of 31 patients treated per protocol. Missing data are denoted by x. Low HL has been previously associated with short survival and poor response to treatment in

Table 1 The real data in the form (CR, HL, FC)

(0, 8.6, 7.8)	(0, 12.9, x)	(0, 11.5, 8.3)	(1, 11.8, 4.1)	(0, 14.7, 5)	(0, 11.9, 6.7)
(0, 9.5, 9.6)	(1, 13.3, 4.2)	(0, 14.1, x)	(0, 11.1, 6)	(1, 12.2, x)	(1, 15, 3.4)
(1, 17.2, 5.1)	(1, 13.9, 3.1)	(1, 13.1, 9.5)	(0, 9.1, 9.7)	(0, 7.2, 9)	(1, 15.4, 6.8)
(0, 11.7, 6.2)	(0, 10.9, 6.18)	(1, 12.1, 7.6)	(1, 13, x)	(0, 12.7, x)	(1, 10.3, 6)
(0, 10.9, 4.1)	(1, 11.9, 5.9)	(0, 13.9, 4.8)	(1, 14.2, 6)	(1, 11.5, 4.4)	(0, 14.5, 4.5)
(0, 11.9, 5.9)					

**Fig. 4** The fitted empirical ROC curve (*dotted curve*) and the proposed estimator of the ROC curve (*solid curve*) for HL (**a**) and FC (**b**)

disseminated disease (Tonini et al. 2011). High FC is examined as a negative predictor of response to treatment in metastatic kidney cancer patients for the first time.

The estimators of the ROC curves for HL (left) and FC (right) as the predictive factors (positive and negative, respectively) are plotted in Fig. 4.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: Proofs of the theorems

Proof (of Theorem 1) Let us fix any $t \in (0, 1)$ for which $R''(s)$ exists for s near t and $R''(s)$ is continuous at $s = t$. Denote

$$T_i = \frac{t - 1 + F(Y_i)}{h_n}, \quad T_{i,m} = \frac{t - 1 + F_m(Y_i)}{h_n}, \quad (16)$$

$i = 1, 2, \dots, n$. The bias of $\hat{R}_{m,n}(t)$ is given by

$$\begin{aligned} b(\hat{R}_{m,n}(t)) &= E_{F,G}[\hat{R}_{m,n}(t)] - R(t) = \frac{1}{n} \sum_{i=1}^n E_{F,G}[\mathcal{K}(T_{i,m})] - R(t) \\ &= E_{F,G}[\mathcal{K}(T_{1,m})] - R(t). \end{aligned}$$

Applying Taylor expansion to function $\mathcal{K}(T_{1,m})$ at T_1 , we obtain

$$\begin{aligned} b(\hat{R}_{m,n}(t)) &= E_{F,G}[\mathcal{K}(T_{1,m})] - R(t) \\ &= E_{F,G}[\mathcal{K}(T_1)] + \sum_{k=1}^{\infty} \frac{1}{k!} E_{F,G}[\mathcal{K}^{(k)}(T_1)(T_{1,m} - T_1)^k] - R(t) \\ &= E_{F,G}[\mathcal{K}(T_1)] + \sum_{k=1}^3 \frac{1}{k!} E_{F,G}[\mathcal{K}^{(k)}(T_1)(T_{1,m} - T_1)^k] - R(t) \\ &:= I_0 + \sum_{k=1}^3 I_k - R(t). \end{aligned} \tag{17}$$

The first term in the sum on the right side of equality (17) is equal to

$$\begin{aligned} I_0 &= E_{F,G}[\mathcal{K}(T_1)] = \int_{-\infty}^{\infty} \mathcal{K}\left(\frac{t-1+F(y)}{h_n}\right) g(y) dy \\ &= \mathcal{K}\left(\frac{t-1+F(y)}{h_n}\right) G(y) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} K\left(\frac{t-1+F(y)}{h_n}\right) \frac{f(y)}{h_n} G(y) dy \\ &= \mathcal{K}\left(\frac{t}{h_n}\right) - \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} K(x) G(F^{-1}(1-t+xh_n)) dx \\ &= \mathcal{K}\left(\frac{t}{h_n}\right) - \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} K(x) [1 - R(t-xh_n)] dx \\ &= \mathcal{K}\left(\frac{t-1}{h_n}\right) + \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} K(x) R(t-xh_n) dx. \end{aligned} \tag{18}$$

Expanding $R(t-xh_n)$ in a Taylor series at t to order 2 and substituting the obtained expansion into (18), we get

$$\begin{aligned} I_0 &= \mathcal{K}\left(\frac{t-1}{h_n}\right) + R(t) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} K(x) dx - R'(t) h_n \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x K(x) dx \\ &\quad + \frac{R''(t)}{2} h_n^2 \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x^2 K(x) dx + o(h_n^2). \end{aligned}$$

Note that for all n greater than some $n_0 \in \mathbb{N}$ we have $\frac{t-1}{h_n} < -1$ and $\frac{t}{h_n} > 1$. Therefore

$$\begin{aligned} I_0 &= R(t) \int_{-1}^1 K(x) dx - R'(t) h_n \int_{-1}^1 x K(x) dx \\ &\quad + \frac{R''(t)}{2} h_n^2 \int_{-1}^1 x^2 K(x) dx + o(h_n^2) = R(t) + \frac{R''(t)}{10} h_n^2 + o(h_n^2). \end{aligned} \quad (19)$$

The expectation I_1 appearing in (17) is equal to zero, because F_m is an unbiased estimator of F , namely

$$\begin{aligned} I_1 &= \frac{1}{1!} E_{F,G} [\mathcal{K}'(T_1) (T_{1,m} - T_1)] \\ &= E_G \left[K \left(\frac{t-1+F(Y_1)}{h_n} \right) E_F \left[\frac{F_m(Y_1) - F(Y_1)}{h_n} \right] \right] = 0. \end{aligned} \quad (20)$$

Applying Lemma 3 to I_2 and I_3 , we obtain

$$I_2 = O\left(\frac{1}{m}\right), \quad I_3 = O\left(\frac{1}{m^2 h_n^2}\right). \quad (21)$$

From (17), (19), (20) i (21), we get

$$b(\hat{R}_{m,n}(t)) = \frac{R''(t)}{10} h_n^2 + O\left(\frac{1}{m} + \frac{1}{m^2 h_n^2}\right) + o(h_n^2).$$

The variance of $\tilde{R}_{m,n}$ is equal to

$$\begin{aligned} \text{Var}(\hat{R}_{m,n}(t)) &= \frac{1}{n^2} \sum_{i=1}^n \text{Var}[\mathcal{K}(T_{i,m})] + \frac{2}{n^2} \sum_{1 \leq i < j \leq n} \text{Cov}[\mathcal{K}(T_{i,m}), \mathcal{K}(T_{j,m})] \\ &= \frac{1}{n} \text{Var}[\mathcal{K}(T_{1,m})] + \frac{n-1}{n} \text{Cov}[\mathcal{K}(T_{1,m}), \mathcal{K}(T_{2,m})] \\ &:= \frac{1}{n} J_1 + \frac{n-1}{n} J_2. \end{aligned} \quad (22)$$

Let us first take care of the variance of $\mathcal{K}(T_{1,m})$, denoted by J_1 .

$$\begin{aligned} J_1 &= E_{F,G} [\mathcal{K}^2(T_{1,m})] - E_{F,G}^2 [\mathcal{K}(T_{1,m})] \\ &= E_{F,G} [\mathcal{K}^2(T_{1,m})] - \left(R(t) + b(\hat{R}_{m,n}(t)) \right)^2 \\ &= E_{F,G} [\mathcal{K}^2(T_{1,m})] - R^2(t) + O\left(h_n^2 + \frac{1}{m} + \frac{1}{m^2 h_n^2} + \frac{1}{m^4 h_n^4}\right). \end{aligned} \quad (23)$$

Applying Taylor expansion to function $\mathcal{K}^2(T_{1,m})$ at T_1 , we obtain

$$\begin{aligned} E_{F,G}[\mathcal{K}^2(T_{1,m})] &= E_{F,G}[\mathcal{K}^2(T_1)] + \sum_{k=1}^{\infty} \frac{1}{k!} E_{F,G}[(\mathcal{K}^2)^{(k)}(T_1)(T_{1,m} - T_1)^k] \\ &= E_{F,G}[\mathcal{K}^2(T_1)] + \sum_{k=1}^6 \frac{1}{k!} E_{F,G}[(\mathcal{K}^2)^{(k)}(T_1)(T_{1,m} - T_1)^k] \\ &:= J_{1,0} + \sum_{k=1}^6 J_{1,k}, \end{aligned}$$

and in consequence, using (23),

$$J_1 = J_{1,0} + \sum_{k=1}^6 J_{1,k} - R^2(t) + O\left(h_n^2 + \frac{1}{m} + \frac{1}{m^2 h_n^2} + \frac{1}{m^4 h_n^4}\right). \quad (24)$$

With respect to $J_{1,0}$, we have

$$\begin{aligned} J_{1,0} &= E_{F,G}[\mathcal{K}^2(T_1)] = E_G\left[\mathcal{K}^2\left(\frac{t-1+F(Y_1)}{h_n}\right)\right] \\ &= \int_{-\infty}^{\infty} \mathcal{K}^2\left(\frac{t-1+F(y)}{h_n}\right) g(y) dy = \mathcal{K}^2\left(\frac{t-1+F(y)}{h_n}\right) G(y) \Big|_{-\infty}^{\infty} \\ &\quad - \int_{-\infty}^{\infty} 2\mathcal{K}\left(\frac{t-1+F(y)}{h_n}\right) K\left(\frac{t-1+F(y)}{h_n}\right) \frac{f(y)}{h_n} G(y) dy \\ &= \mathcal{K}^2\left(\frac{t}{h_n}\right) - 2 \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \mathcal{K}(x) K(x) G\left(F^{-1}(1-t+xh_n)\right) dx \\ &= \mathcal{K}^2\left(\frac{t-1}{h_n}\right) + 2 \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \mathcal{K}(x) K(x) R(t-xh_n) dx. \end{aligned}$$

Expanding $R(t-xh_n)$ in a Taylor series at t to order 2, we get

$$\begin{aligned} J_{1,0} &= \mathcal{K}^2\left(\frac{t-1}{h_n}\right) + 2R(t) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \mathcal{K}(x) K(x) dx \\ &\quad - 2R'(t)h_n \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x \mathcal{K}(x) K(x) dx + o(h_n) \\ &= 2R(t) \int_{-1}^1 \mathcal{K}(x) K(x) dx - 2R'(t)h_n \int_{-1}^1 x \mathcal{K}(x) K(x) dx + o(h_n) \\ &= R(t) - \frac{9}{35} R'(t)h_n + o(h_n). \end{aligned} \quad (25)$$

Let us now return to equation (24). Using the fact that F_m is an unbiased estimator of F , we get $J_{1,1} = 0$. Therefore, applying Lemma 3 to $J_{1,k}$, $k = 2, 3, \dots, 6$, we obtain

$$\sum_{k=1}^6 J_{1,k} = O\left(\sum_{k=2}^6 \frac{m^{\lfloor \frac{k}{2} \rfloor}}{m^k h_n^{k-1}}\right) = O\left(\frac{1}{m h_n} + \frac{1}{m^2 h_n^3} + \frac{1}{m^3 h_n^5}\right),$$

which in combination with (24) and (25), gives

$$J_1 = R(t)(1 - R(t)) - \frac{9}{35} R'(t) h_n + O\left(h_n^2 + \frac{1}{m h_n} + \frac{1}{m^2 h_n^3} + \frac{1}{m^3 h_n^5}\right) \quad (26)$$

With respect to the covariance $\text{Cov}[\mathcal{K}(T_{1,m}), \mathcal{K}(T_{2,m})]$ denoted in (22) by J_2 , using the fact that $T_{1,m}$ and $T_{2,m}$ have the same distributions, we can write

$$\begin{aligned} J_2 &= E_{F,G}[\mathcal{K}(T_{1,m}) \mathcal{K}(T_{2,m})] - E_{F,G}[\mathcal{K}(T_{1,m})] E_{F,G}[\mathcal{K}(T_{2,m})] = \\ &= E_{F,G}[\mathcal{K}(T_{1,m}) \mathcal{K}(T_{2,m})] - E_{F,G}^2[\mathcal{K}(T_{1,m})]. \end{aligned}$$

Using now the fact that T_1 and T_2 are i.i.d., we get

$$\begin{aligned} J_2 &= E_{F,G}[\mathcal{K}(T_{1,m}) \mathcal{K}(T_{2,m})] - E_G^2[\mathcal{K}(T_1)] + E_G^2[\mathcal{K}(T_1)] - E_{F,G}^2[\mathcal{K}(T_{1,m})] \\ &= E_{F,G}[\mathcal{K}(T_{1,m}) \mathcal{K}(T_{2,m})] - E_G[\mathcal{K}(T_1) \mathcal{K}(T_2)] \\ &\quad + (E_G^2[\mathcal{K}(T_1)] - E_{F,G}^2[\mathcal{K}(T_{1,m})]). \end{aligned}$$

Applying the two-variable Taylor formula to $\mathcal{K}(T_{1,m}) \mathcal{K}(T_{2,m})$ and expanding the function at (T_1, T_2) , we obtain

$$\begin{aligned} J_2 &= \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{i=0}^k \binom{k}{i} E_{F,G}[\mathcal{K}^{(i)}(T_1) \mathcal{K}^{(k-i)}(T_2) (T_{1,m} - T_1)^i (T_{2,m} - T_2)^{k-i}] \\ &\quad + (E_G^2[\mathcal{K}(T_1)] - E_{F,G}^2[\mathcal{K}(T_{1,m})]) \\ &= \sum_{k=1}^6 \frac{1}{k!} \sum_{i \in \mathcal{J}_k} \binom{k}{i} E_{F,G}[\mathcal{K}^{(i)}(T_1) \mathcal{K}^{(k-i)}(T_2) (T_{1,m} - T_1)^i (T_{2,m} - T_2)^{k-i}] \\ &\quad + (E_G^2[\mathcal{K}(T_1)] - E_{F,G}^2[\mathcal{K}(T_{1,m})]), \end{aligned} \quad (27)$$

where $\mathcal{J}_k = \{i \in \mathbb{N} : 0 \vee (k-3) \leq i \leq k \wedge 3\}$. Let us now take care of the last term of the sum in (27). It is easy to check that

$$\begin{aligned} E_G^2[\mathcal{K}(T_1)] - E_{F,G}^2[\mathcal{K}(T_{1,m})] &= -2E_G[\mathcal{K}(T_1)] (E_{F,G}[\mathcal{K}(T_{1,m})] - E_G[\mathcal{K}(T_1)]) \\ &\quad - (E_{F,G}[\mathcal{K}(T_{1,m})] - E_G[\mathcal{K}(T_1)])^2. \end{aligned}$$

Form (17), (20) and (21), we know that $E_{F,G} [\mathcal{K} (T_{1,m})] - E_G [\mathcal{K} (T_1)]$ is equal to

$$\sum_{k=1}^3 \frac{1}{k!} E_{F,G} [\mathcal{K}^{(k)} (T_1) (T_{1,m} - T_1)^k] = O \left(\frac{1}{m} + \frac{1}{m^2 h_n^2} \right)$$

and in result

$$\begin{aligned} E_G^2 [\mathcal{K} (T_1)] - E_{F,G}^2 [\mathcal{K} (T_{1,m})] \\ = -2 E_G [\mathcal{K} (T_1)] \sum_{k=1}^3 \frac{1}{k!} E_{F,G} [\mathcal{K}^{(k)} (T_1) (T_{1,m} - T_1)^k] \\ + O \left(\frac{1}{m^2} + \frac{1}{m^3 h_n^2} + \frac{1}{m^4 h_n^4} \right). \end{aligned} \quad (28)$$

Again using the fact that T_1 and T_2 , and $T_{1,m}$ and $T_{2,m}$ have the same distributions, and T_1 and T_2 are independent, equality (28) may be written in the following form

$$\begin{aligned} E_G^2 [\mathcal{K} (T_1)] - E_{F,G}^2 [\mathcal{K} (T_{1,m})] \\ = -2 \sum_{k=1}^3 \frac{1}{k!} E_{F,G} [\mathcal{K}^{(k)} (T_1) \mathcal{K} (T_2) (T_{1,m} - T_1)^k] + O \left(\frac{1}{m^2} + \frac{1}{m^3 h_n^2} + \frac{1}{m^4 h_n^4} \right) \\ = - \sum_{k=1}^3 \frac{1}{k!} E_{F,G} [\mathcal{K}^{(k)} (T_1) \mathcal{K} (T_2) (T_{1,m} - T_1)^k] \\ - \sum_{k=1}^3 \frac{1}{k!} E_{F,G} [\mathcal{K} (T_1) \mathcal{K}^{(k)} (T_2) (T_{2,m} - T_2)^k] + O \left(\frac{1}{m^2} + \frac{1}{m^3 h_n^2} + \frac{1}{m^4 h_n^4} \right). \end{aligned}$$

Therefore, using (27), we get

$$\begin{aligned} J_2 = \sum_{k=2}^6 \frac{1}{k!} \sum_{j \in \mathcal{J}_k} \binom{k}{j} E_{F,G} [\mathcal{K}^{(j)} (T_1) \mathcal{K}^{(k-j)} (T_2) (T_{1,m} - T_1)^j (T_{2,m} - T_2)^{k-j}] \\ + O \left(\frac{1}{m^2} + \frac{1}{m^3 h_n^2} + \frac{1}{m^4 h_n^4} \right) := \sum_{k=2}^6 J_{2,k} + O \left(\frac{1}{m^2} + \frac{1}{m^3 h_n^2} + \frac{1}{m^4 h_n^4} \right). \end{aligned} \quad (29)$$

where $\mathcal{J}_k = \{j \in \mathbb{N} : 1 \vee (k-3) \leq j \leq (k-1) \wedge 3\}$. The first term in the sum on the right side of equality (29) is equal to

$$\begin{aligned} J_{2,2} &= \frac{1}{2!} \binom{2}{1} E_{F,G} \left[\mathcal{K}'(T_1) \mathcal{K}'(T_2) (T_{1,m} - T_1)(T_{2,m} - T_2) \right] \\ &= \frac{1}{h_n^2} E_G [K(T_1) K(T_2) E_F[F_m(Y_1) - F(Y_1)][F_m(Y_2) - F(Y_2)]] \\ &= \frac{1}{mh_n^2} E_G [K(T_1) K(T_2) H_{1,1}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2))], \end{aligned}$$

where

$$H_{1,1}(x, y, z) = x - yz = x^1 y^0 z^0 - x^0 y^1 z^1.$$

Applying Lemma 4 to $J_{2,2}$ expressed as above, we obtain

$$J_{2,2} = \frac{1}{mh_n^2} \left(h_n^2 [R'(t)]^2 t(1-t) + h_n^3 \beta [R'(t)]^2 + o(h_n^3) \right),$$

where

$$\beta = \int_{-1}^1 K(u) \int_{-1}^u x K(x) dx du + \int_{-1}^1 u K(u) \int_u^1 K(x) dx du = -\frac{9}{35}.$$

Simplifying, we get

$$J_{2,2} = \frac{t(1-t)[R'(t)]^2}{m} - \frac{\frac{9}{35}[R'(t)]^2 h_n}{m} + o\left(\frac{h_n}{m}\right). \quad (30)$$

The term $J_{2,4}$ is the sum of three another terms from which the first one and the third one have the same values, and their sum is equal to

$$\begin{aligned} J_{2,4,1} + J_{2,4,3} &= 2J_{2,4,1} = \frac{2}{4!} \binom{4}{1} E_{F,G} \left[\mathcal{K}'(T_1) \mathcal{K}^{(3)}(T_2) (T_{1,m} - T_1)(T_{2,m} - T_2)^3 \right] \\ &= \frac{1}{3h_n^4} E_G \left[K(T_1) K^{(2)}(T_2) E_F[F_m(Y_1) - F(Y_1)][F_m(Y_2) - F(Y_2)]^3 \right], \end{aligned}$$

and the second one is equal to

$$\begin{aligned} J_{2,4,2} &= \frac{1}{4!} \binom{4}{2} E_{F,G} \left[\mathcal{K}^{(2)}(T_1) \mathcal{K}^{(2)}(T_2) (T_{1,m} - T_1)^2 (T_{2,m} - T_2)^2 \right] \\ &= \frac{1}{4h_n^4} E_G \left[K'(T_1) K'(T_2) E_F[F_m(Y_1) - F(Y_1)]^2 [F_m(Y_2) - F(Y_2)]^2 \right]. \end{aligned}$$

One can check that

$$\begin{aligned} E_F[F_m(Y_1) - F(Y_1)][F_m(Y_2) - F(Y_2)]^3 \\ &= \frac{H_{1,3}^{(1)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2))}{m^2} + \frac{H_{1,3}^{(2)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2))}{m^3}, \\ E_F[F_m(Y_1) - F(Y_1)]^2[F_m(Y_2) - F(Y_2)]^2 \\ &= \frac{H_{2,2}^{(1)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2))}{m^2} + \frac{H_{2,2}^{(2)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2))}{m^3}, \end{aligned}$$

where

$$\begin{aligned} H_{1,3}^{(1)}(x, y, z) &= 3xz - 3yz^2 - 3xz^2 + 3yz^3, \\ H_{1,3}^{(2)}(x, y, z) &= x - yz - 6xz + 6yz^2 + 6xz^2 - 6yz^3, \\ H_{2,2}^{(1)}(x, y, z) &= yz - yz^2 - y^2z + 2x^2 - 4xyz + 3y^2z^2, \\ H_{2,2}^{(2)}(x, y, z) &= x - yz - 2xz - 2xy + 2yz^2 + 2y^2z - 2x^2 + 8xyz - 6y^2z^2. \end{aligned}$$

Therefore

$$\begin{aligned} J_{2,4} &= J_{2,4,1} + J_{2,4,2} + J_{2,4,3} \\ &= \frac{1}{3m^2h_n^4} E_G \left[K(T_1) K^{(2)}(T_2) H_{1,3}^{(1)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2)) \right] \\ &\quad + \frac{1}{3m^3h_n^4} E_G \left[K(T_1) K^{(2)}(T_2) H_{1,3}^{(2)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2)) \right] \\ &\quad + \frac{1}{4m^2h_n^4} E_G \left[K'(T_1) K'(T_2) H_{2,2}^{(1)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2)) \right] \\ &\quad + \frac{1}{4m^3h_n^4} E_G \left[K'(T_1) K'(T_2) H_{2,2}^{(2)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2)) \right]. \quad (31) \end{aligned}$$

Applying Lemma 4 to each of the expectations appearing in (31), we get

$$J_{2,4} = \frac{1}{3m^2h_n^4} \left(h_n^2 \gamma [R'(t)]^2 3t^2(1-t)^2 + O(h_n^3) \right) + O\left(\frac{h_n^2 + mh_n^3 + h_n^3}{m^3h_n^4} \right),$$

where

$$\gamma = \int_{-1}^1 K(x) dx \int_{-1}^1 K^{(2)}(x) dx = -3.$$

After simplification, we have

$$J_{2,4} = \frac{-3t^2(1-t)^2[R'(t)]^2}{m^2h_n^2} + O\left(\frac{1}{m^2h_n} + \frac{1}{m^3h_n^2} \right). \quad (32)$$

The last term in the sum on the right side of equality (29) is equal to

$$\begin{aligned} J_{2,6} &= \frac{1}{6!} \binom{6}{3} E_{F,G} \left[\mathcal{K}^{(3)}(T_1) \mathcal{K}^{(3)}(T_2) (T_{1,m} - T_1)^3 (T_{2,m} - T_2)^3 \right] \\ &= \frac{1}{36h_n^6} E_G \left[K^{(2)}(T_1) K^{(2)}(T_2) E_F [F_m(Y_1) - F(Y_1)]^3 [F_m(Y_2) - F(Y_2)]^3 \right]. \end{aligned} \quad (33)$$

Again, one can check that

$$\begin{aligned} E_F [F_m(Y_1) - F(Y_1)]^3 [F_m(Y_2) - F(Y_2)]^3 &= \frac{H_{3,3}^{(1)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2))}{m^3} \\ &+ \frac{H_{3,3}^{(2)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2))}{m^4} + \frac{H_{3,3}^{(3)}(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2))}{m^5}, \end{aligned}$$

where

$$\begin{aligned} H_{3,3}^{(1)}(x, y, z) &= 9xyz - 9y^2z^2 - 9xyz^2 - 9xy^2z + 9y^2z^3 + 9y^3z^2 \\ &\quad + 27xy^2z^2 - 18x^2yz + 6x^3 - 15y^3z^3, \\ H_{3,3}^{(2)}(x, y, z) &= yz + 3xy + 3xz - 6yz^2 - 6y^2z - 3xy^2 \\ &\quad - 3xz^2 + 5yz^3 + 5y^3z + 18x^2, \\ &\quad - 72xyz + 63y^2z^2 - 36x^2y - 36x^2z + 108xyz^2 + 108xy^2z \\ &\quad - 78y^2z^3 - 78y^3z^2 - 234xy^2z^2 + 126x^2yz - 18x^3 + 130y^3z^3 \end{aligned}$$

and

$$\begin{aligned} H_{3,3}^{(3)}(x, y, z) &= x - yz - 6xy - 6xz + 6yz^2 + 6y^2z + 6xy^2 + 6xz^2 - 6yz^3 - 6y^3z \\ &\quad - 18x^2 + 72xyz - 54y^2z^2 + 36x^2y + 36x^2z - 108xyz^2 - 108xy^2z \\ &\quad + 72y^2z^3 + 72y^3z^2 + 216xy^2z^2 - 108x^2yz + 12x^3 - 120y^3z^3. \end{aligned}$$

Applying now Lemma 4 to the expectation in (33), we obtain

$$J_{2,6} = \frac{1}{36m^3h_n^6} \left(h_n^2 \delta [R'(t)]^2 15t^3 (1-t)^3 + O(h_n^3) \right) + O \left(\frac{mh_n^2 + h_n^2}{m^5h_n^6} \right),$$

where

$$\delta = \left(\int_{-1}^1 K^{(2)}(x) dx \right)^2 = 9.$$

Simplifying, we get

$$J_{2,6} = \frac{\frac{15}{4}t^3(1-t)^3[R'(t)]^2}{m^3h_n^4} + O \left(\frac{1}{m^3h_n^3} + \frac{1}{m^4h_n^4} \right). \quad (34)$$

The remaining terms $J_{2,3}$ and $J_{2,5}$ in the sum on the right side of equality (29), we can estimate using Lemma 3

$$J_{2,3} = O\left(\frac{1}{m^2 h_n}\right), \quad J_{2,5} = O\left(\frac{1}{m^3 h_n^3}\right). \quad (35)$$

Combining now (29), (30), (32), (34) and (35), we obtain

$$\begin{aligned} J_2 &= \frac{t(1-t)[R'(t)]^2}{m} - \frac{\frac{9}{35}[R'(t)]^2 h_n}{m} - \frac{3t^2(1-t)^2[R'(t)]^2}{m^2 h_n^2} \\ &\quad + \frac{\frac{15}{4}t^3(1-t)^3[R'(t)]^2}{m^3 h_n^4} \\ &\quad + O\left(\frac{1}{m^2 h_n} + \frac{1}{m^3 h_n^3} + \frac{1}{m^4 h_n^4}\right) + o\left(\frac{h_n}{m}\right). \end{aligned} \quad (36)$$

Finally, substituting (26) and (36) into (22), after simplification, we have

$$\begin{aligned} \text{Var}(\hat{R}_{m,n}(t)) &= \frac{R(t)(1-R(t))}{n} + \frac{t(1-t)[R'(t)]^2}{m} - \frac{\frac{9}{35}R'(t)\left[\frac{m}{n} - R'(t)\right]h_n}{m} \\ &\quad - \frac{3t^2(1-t)^2[R'(t)]^2}{m^2 h_n^2} + \frac{\frac{15}{4}t^3(1-t)^3[R'(t)]^2}{m^3 h_n^4} \\ &\quad + O\left(\frac{m+n}{m^2 n h_n} + \frac{m+n}{m^3 n h_n^3} + \frac{m+n h_n}{m^4 n h_n^5}\right) + o\left(\frac{h_n}{m}\right), \end{aligned}$$

which completes the proof. \square

Proof (of Theorem 2) From (11) and Theorem 1, under the assumption $nh_n^2 \rightarrow \infty$, we can write

$$\text{MSE}(\tilde{R}_n^*(t)) = \frac{A(t)}{n} + O\left(\frac{\lambda - \lambda_n}{n}\right) + O\left(n^{-\frac{4}{3}}\right) \quad (37)$$

and

$$\text{MSE}(\hat{R}_n(t)) = \frac{A(t)}{n} - \frac{B(t)}{n^2 h_n^2} + O\left(\frac{\lambda - \lambda_n}{n}\right) + O\left(\frac{h_n}{n} + h_n^4\right) + o\left(\frac{1}{n^2 h_n^2}\right), \quad (38)$$

where

$$\begin{aligned} A(t) &= \frac{R(t)[1-R(t)]\lambda + t(1-t)[R'(t)]^2}{\lambda} > 0, \\ B(t) &= \frac{3t^2(1-t)^2[R'(t)]^2}{\lambda^2} \geq 0. \end{aligned} \quad (39)$$

From the definition (8) of $b_n = b_n(t)$, we have

$$\text{MSE}(\tilde{R}_{b_n}^*(t)) \leq \text{MSE}(\hat{R}_n(t)), \quad (40)$$

$$\text{MSE}(\tilde{R}_{b_n-1}^*(t)) > \text{MSE}(\hat{R}_n(t)). \quad (41)$$

Substituting (37) and (38) into (40), we get

$$\begin{aligned} \frac{A(t)}{b_n} + O\left(\frac{\lambda - \lambda_{b_n}}{b_n}\right) + O\left(b_n^{-\frac{4}{3}}\right) &\leq \frac{A(t)}{n} - \frac{B(t)}{n^2 h_n^2} + O\left(\frac{\lambda - \lambda_n}{n}\right) \\ &\quad + O\left(\frac{h_n}{n} + h_n^4\right) + o\left(\frac{1}{n^2 h_n^2}\right) \end{aligned}$$

and after some simple transformations, we obtain

$$\frac{b_n}{n} \geq \frac{1 + O(\lambda - \lambda_{b_n}) + O(b_n^{-\frac{1}{3}})}{1 - \frac{B(t)}{A(t)} \frac{1}{n h_n^2} + O(\lambda - \lambda_n) + O(h_n + n h_n^4) + o\left(\frac{1}{n h_n^2}\right)}. \quad (42)$$

From (37), (38) and (41), after analogical transformations, we get

$$\frac{b_n}{n} < \frac{1 + O(\lambda - \lambda_{b_n-1}) + O((b_n - 1)^{-\frac{1}{3}})}{1 - \frac{B(t)}{A(t)} \frac{1}{n h_n^2} + O(\lambda - \lambda_n) + O(h_n + n h_n^4) + o\left(\frac{1}{n h_n^2}\right)} + \frac{1}{n}. \quad (43)$$

Combining (42) and (43), and using the assumptions $h_n \rightarrow 0$, $n h_n^2 \rightarrow \infty$, $n h_n^4 \rightarrow 0$, $\lambda_n \rightarrow \lambda$ and the fact that $b_n \rightarrow \infty$, we obtain

$$\lim_{n \rightarrow \infty} \frac{b_n}{n} = 1. \quad (44)$$

Suppose now that $n h_n^3 \rightarrow 0$ and $\lambda - \lambda_n = O(n^{-1/3})$. Subtracting 1 from both sides of inequalities (42) and (43), and multiplying them by $n h_n^2$, we get

$$h_n^2(b_n - n) \geq \frac{\frac{B(t)}{A(t)} + O\left\{\left(1 + \left(\frac{b_n}{n}\right)^{-\frac{1}{3}}\right) n^{\frac{2}{3}} h_n^2\right\} + O(n h_n^3 + n^2 h_n^6) + o(1)}{1 - \frac{B(t)}{A(t)} \frac{1}{n h_n^2} + O(n^{-\frac{1}{3}}) + O(h_n + n h_n^4) + o\left(\frac{1}{n h_n^2}\right)} \quad (45)$$

and

$$h_n^2(b_n - n) < \frac{\frac{B(t)}{A(t)} + O\left\{\left(1 + \left(\frac{b_n-1}{n}\right)^{-\frac{1}{3}}\right) n^{\frac{2}{3}} h_n^2\right\} + O(n h_n^3 + n^2 h_n^6) + o(1)}{1 - \frac{B(t)}{A(t)} \frac{1}{n h_n^2} + O(n^{-\frac{1}{3}}) + O(h_n + n h_n^4) + o\left(\frac{1}{n h_n^2}\right)}, \quad (46)$$

respectively. Using the assumptions $h_n \rightarrow 0$, $nh_n^2 \rightarrow \infty$, $nh_n^3 \rightarrow 0$ and equality (44), from inequalities (45) and (46), we obtain

$$\lim_{n \rightarrow \infty} h_n^2(b_n - n) = \frac{B(t)}{A(t)} = \frac{3t^2(1-t)^2[R'(t)]^2}{R(t)[1-R(t)]\lambda^2 + t(1-t)[R'(t)]^2\lambda} \geq 0.$$

Suppose now that $nh_n^2 \rightarrow \delta \in (0, \infty)$. Then, from Theorem 1, we have

$$\text{MSE}(\hat{R}_n(t)) = \frac{A(t)}{n} - \frac{B(t)}{n^2 h_n^2} + \frac{C(t)}{n^3 h_n^4} + O\left(\frac{\lambda - \lambda_n}{n}\right) + O\left(n^{-\frac{3}{2}}\right), \quad (47)$$

where $A(t)$ and $B(t)$ are given by (39), and

$$C(t) = \frac{\frac{15}{4}t^3(1-t)^3[R'(t)]^2}{\lambda^3} \geq 0. \quad (48)$$

Substituting (37) and (47) into (40), after some transformations, we get

$$\frac{b_n}{n} \geq \frac{1 + O(\lambda - \lambda_{b_n}) + O\left(b_n^{-\frac{1}{3}}\right)}{1 - \frac{B(t)}{A(t)} \frac{1}{nh_n^2} + \frac{C(t)}{A(t)} \frac{1}{n^2 h_n^4} + O(\lambda - \lambda_n) + O\left(n^{-\frac{1}{2}}\right)}. \quad (49)$$

Analogously, substituting (37) and (47) into (41), we obtain

$$\frac{b_n}{n} < \frac{1 + O(\lambda - \lambda_{b_n-1}) + O\left((b_n - 1)^{-\frac{1}{3}}\right)}{1 - \frac{B(t)}{A(t)} \frac{1}{nh_n^2} + \frac{C(t)}{A(t)} \frac{1}{n^2 h_n^4} + O(\lambda - \lambda_n) + O\left(n^{-\frac{1}{2}}\right)} + \frac{1}{n}. \quad (50)$$

Combining (49) and (50), and using the condition $nh_n^2 \rightarrow \delta \in (0, \infty)$, we get

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{b_n}{n} &= \frac{1}{1 - \frac{B(t)}{A(t)} \frac{1}{\delta} + \frac{C(t)}{A(t)} \frac{1}{\delta^2}} = \frac{\delta^2}{\delta^2 - \frac{B(t)\delta - C(t)}{A(t)}} \\ &= \frac{\delta^2}{\delta^2 - \frac{3t^2(1-t)^2[R'(t)]^2\left\{\lambda\delta - \frac{5}{4}t(1-t)\right\}}{R(t)[1-R(t)]\lambda^3 + t(1-t)[R'(t)]^2\lambda^2}}. \end{aligned}$$

One can easily see that for $\delta > \frac{5}{4\lambda}t(1-t)$, the above limit is greater than 1, which completes the proof. \square

Appendix 2: Some useful lemmas

Lemma 1 (Znidaric 2009, Corollary 1) *Let $\mu_k(m, p)$ denote the k -th central moment of the binomial distribution $\mathcal{B}(m, p)$, i.e. $\mu_k(m, p) = E(X - mp)^k$, where $X \sim \mathcal{B}(m, p)$. Then*

- (1) $\mu_k(m, p)$ is a polynomial of degree $\lfloor \frac{k}{2} \rfloor$ in m ,
- (2) $\mu_k(m, p)$ is a polynomial of degree k in p .

Lemma 2 Let $\mu_{k,l}(m, \mathbf{p})$ denote the mixed central moment of order $k + l$ of the multinomial distribution $\mathcal{M}_3(m, \mathbf{p})$, i.e. $\mu_{k,l}(m, \mathbf{p}) = E(X_1 - mp_1)^k (X_2 - mp_2)^l$, where $\mathbf{X} = (X_1, X_2, X_3) \sim \mathcal{M}_3(m, \mathbf{p})$, $\mathbf{p} = (p_1, p_2, p_3)$, $p_1 + p_2 + p_3 = 1$. Then

- (1) $\mu_{k,l}(m, \mathbf{p})$ is a polynomial of degree at most $\lfloor \frac{k+l}{2} \rfloor$ in m .
- (2) $\mu_{k,l}(m, \mathbf{p})$ is a polynomial of degree at most $k + l$ in p_1 and p_2 .

Proof Let $k, l \in \mathbb{N}$ and let $\mathcal{X} = \{(x_1, x_2, x_3) \in \mathbb{N}^3 : \sum_{i=1}^3 x_i = m\}$.

$$\begin{aligned} \mu_{k,l} &:= \mu_{k,l}(m, \mathbf{p}) = E(X_1 - mp_1)^k (X_2 - mp_2)^l \\ &= \sum_{\mathcal{X}} P(X_1 = x_1, X_2 = x_2, X_3 = x_3) (x_1 - mp_1)^k (x_2 - mp_2)^l \\ &= \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{x_3} (x_1 - mp_1)^k (x_2 - mp_2)^l. \end{aligned}$$

Differentiating $\mu_{k,l}$ with respect to p_1 , we get

$$\begin{aligned} \frac{\partial \mu_{k,l}}{\partial p_1} &= \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} x_1 p_1^{x_1-1} p_2^{x_2} (1 - p_1 - p_2)^{x_3} (x_1 - mp_1)^k (x_2 - mp_2)^l \\ &\quad - \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} x_3 p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{x_3-1} (x_1 - mp_1)^k (x_2 - mp_2)^l \\ &\quad - mk \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{x_3} (x_1 - mp_1)^{k-1} (x_2 - mp_2)^l. \end{aligned} \quad (51)$$

Analogously, differentiating $\mu_{k,l}$ with respect to p_2 , we obtain

$$\begin{aligned} \frac{\partial \mu_{k,l}}{\partial p_2} &= \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} x_2 p_1^{x_1} p_2^{x_2-1} (1 - p_1 - p_2)^{x_3} (x_1 - mp_1)^k (x_2 - mp_2)^l \\ &\quad - \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} x_3 p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{x_3-1} (x_1 - mp_1)^k (x_2 - mp_2)^l \\ &\quad - ml \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{x_3} (x_1 - mp_1)^k (x_2 - mp_2)^{l-1}. \end{aligned} \quad (52)$$

Note that the last of the three terms appearing in (51) and (52) are equal $-mk\mu_{k-1,l}$ and $-ml\mu_{k,l-1}$, respectively. Hence

$$\left[\frac{\partial \mu_{k,l}}{\partial p_1} + mk\mu_{k-1,l} \right] p_1(1 - p_1 - p_2) + \left[\frac{\partial \mu_{k,l}}{\partial p_2} + ml\mu_{k,l-1} \right] p_2(1 - p_1 - p_2)$$

$$\begin{aligned}
&= \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} x_1^{x_1} p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{x_3+1} (x_1-mp_1)^k (x_2-mp_2)^l \\
&\quad - \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} x_3 p_1^{x_1+1} p_2^{x_2} (1-p_1-p_2)^{x_3} (x_1-mp_1)^k (x_2-mp_2)^l \\
&\quad + \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} x_2 p_1^{x_1} p_2^{x_2+1} (1-p_1-p_2)^{x_3+1} (x_1-mp_1)^k (x_2-mp_2)^l \\
&\quad - \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} x_3 p_1^{x_1} p_2^{x_2+1} (1-p_1-p_2)^{x_3} (x_1-mp_1)^k (x_2-mp_2)^l \\
&= \sum_{\mathcal{X}} [x_1(1-p_1-p_2) - x_3 p_1 + x_2(1-p_1-p_2) - x_3 p_2] \\
&\quad \times \frac{m!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{x_3} (x_1-mp_1)^k (x_2-mp_2)^l. \tag{53}
\end{aligned}$$

Using the fact that $x_3 = m - x_1 - x_2$, we get

$$x_1(1-p_1-p_2) - x_3 p_1 + x_2(1-p_1-p_2) - x_3 p_2 = (x_1-mp_1) + (x_2-mp_2),$$

which in combination with (53), gives

$$\begin{aligned}
&\left[\frac{\partial \mu_{k,l}}{\partial p_1} + mk\mu_{k-1,l} \right] p_1(1-p_1-p_2) + \left[\frac{\partial \mu_{k,l}}{\partial p_2} + ml\mu_{k,l-1} \right] p_2(1-p_1-p_2) \\
&= \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{x_3} (x_1-mp_1)^{k+1} (x_2-mp_2)^l \\
&\quad + \sum_{\mathcal{X}} \frac{m!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{x_3} (x_1-mp_1)^k (x_2-mp_2)^{l+1} \\
&= \mu_{k+1,l} + \mu_{k,l+1}. \tag{54}
\end{aligned}$$

The recursive formula (54) and the initial conditions

$$\begin{aligned}
\mu_{1,0} &= \mu_{0,1} = 0, \\
\mu_{2,0} &= mp_1(1-p_1), \mu_{1,1} = -mp_1 p_2, \mu_{0,2} = mp_2(1-p_2), \tag{55}
\end{aligned}$$

let us to prove the lemma, using mathematical induction over $s = k + l$. We give the proof only of the first assertion of the lemma. The proof of the second is essentially the same. Equalities (55) indicate that, the thesis is true for $s = 1$ and $s = 2$. Suppose that the thesis is satisfied for some natural $s_0 \geq 2$ and $s_0 + 1$, i.e. for any $k, l \in \mathbb{N}$ such that $k + l = s_0$ or $k + l = s_0 + 1$, $\mu_{k,l}$ is a polynomial (of variable m) of degree at most $\lfloor \frac{k+l}{2} \rfloor$. We need to show that all $\mu_{k,l}$, where $(k, l) \in \{(s_0 + 2, 0), (s_0 + 1, 1), (s_0, 2), \dots, (1, s_0 + 1), (0, s_0 + 2)\}$, are polynomials of degree at most $\lfloor \frac{s_0+2}{2} \rfloor$.

It follows from Lemma 1 that $\mu_{s_0+2,0} = E(X_1 - mp_1)^{s_0+2}$ is a polynomial of degree $\lfloor \frac{s_0+2}{2} \rfloor$. Assume that $\mu_{s_0+2,0}, \mu_{s_0+1,1}, \mu_{s_0,2}, \dots, \mu_{s_0+2-s_1,s_1}$, where $0 \leq s_1 <$

$s_0 + 2$, satisfied the thesis of the lemma. We show that then also $\mu_{s_0+1-s_1, s_1+1}$ satisfied the thesis. Indeed, from (54) (taking $k = s_0 + 1 - s_1, l = s_1$), we have

$$\begin{aligned} \mu_{s_0+1-s_1, s_1+1} &= -\mu_{s_0+2-s_1, s_1} \\ &+ \left[\frac{\partial \mu_{s_0+1-s_1, s_1}}{\partial p_1} + m(s_0 + 1 - s_1) \mu_{s_0-s_1, s_1} \right] p_1(1 - p_1 - p_2) \\ &+ \left[\frac{\partial \mu_{s_0+1-s_1, s_1}}{\partial p_2} + m s_1 \mu_{s_0+1-s_1, s_1-1} \right] p_2(1 - p_1 - p_2). \end{aligned} \quad (56)$$

By the induction hypothesis, $\mu_{s_0+1-s_1, s_1}$ is a polynomial of degree at most $\lfloor \frac{s_0+1}{2} \rfloor$ and $\mu_{s_0-s_1, s_1}$ and $\mu_{s_0+1-s_1, s_1-1}$ are polynomials of degree at most $\lfloor \frac{s_0}{2} \rfloor$. Moreover, as we assumed, $\mu_{s_0+2-s_1, s_1}$ is a polynomial of degree at most $\lfloor \frac{s_0+2}{2} \rfloor$. Thus, using (56), we conclude that also $\mu_{s_0+1-s_1, s_1+1}$ is a polynomial of degree at most $\lfloor \frac{s_0+2}{2} \rfloor$. Hence, one can easily deduce that the degrees of all $\mu_{s_0+2-i, i}$, $i = 0, 1, \dots, s_0 + 2$ are equal at most $\lfloor \frac{s_0+2}{2} \rfloor$, which completes the proof. \square

Lemma 3 Let $\mathbf{X}_m = (X_1, \dots, X_m)$ be a simple sample from a continuous distribution function F and let F_m denote its empirical distribution function. Let Y_1 and Y_2 be independent random variables from a continuous distribution function G and let $R(s) = 1 - G(F^{-1}(1 - s))$. Assume that $R''(s)$ exists for s near $t \in (0, 1)$, $R''(s)$ is continuous at $s = t$, and let ϕ_1 and ϕ_2 be continuous functions supported on $[-1, 1]$. Then for any $h_n > 0$ such that $h_n \rightarrow 0$ and for any $k, n \in \mathbb{N}$, we have:

$$\begin{aligned} (1) \quad E_{F, G} \left[\phi_1(T_1) (T_{1,m} - T_1)^k \right] &= \begin{cases} O \left(\frac{m^{\lfloor \frac{k}{2} \rfloor}}{m^k h_n^{k-2}} \right), & \text{if } \phi_1 \text{ is odd function,} \\ O \left(\frac{m^{\lfloor \frac{k}{2} \rfloor}}{m^k h_n^{k-1}} \right), & \text{otherwise,} \end{cases} \\ (2) \quad E_{F, G} \left[\phi_1(T_1) \phi_2(T_2) (T_{1,m} - T_1)^k (T_{2,m} - T_2)^l \right] &= O \left(\frac{m^{\lfloor \frac{k+l}{2} \rfloor}}{m^{k+l} h_n^{k+l-2}} \right), \end{aligned}$$

where $T_i = \frac{t-1+F(Y_i)}{h_n}$, $T_{i,m} = \frac{t-1+F_m(Y_i)}{h_n}$, $i = 1, 2$.

Proof From the fact that the random variable $m F_m(y) = m F_m(y, \mathbf{X}_m)$ has a binomial distribution $\mathcal{B}(m, F(y))$ for any $y \in \mathbb{R}$, using Lemma 1, we can write

$$E_F[F_m(y) - F(y)]^k = \frac{1}{m^k} E_F[m F_m(y) - m F(y)]^k = \frac{1}{m^k} \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \omega_i(F(y)) m^i,$$

where $\omega_i(x)$, $i = 1, 2, \dots, \lfloor \frac{k}{2} \rfloor$, are polynomials of degree at most k . Hence

$$\begin{aligned}
 E_{F,G} \left[\phi_1(T_1) (T_{1,m} - T_1)^k \right] &= E_{F,G} \left[\phi_1 \left(\frac{t-1+F(Y_1)}{h_n} \right) \frac{[F_m(Y_1) - F(Y_1)]^k}{h_n^k} \right] \\
 &= E_G \left[\phi_1 \left(\frac{t-1+F(Y_1)}{h_n} \right) \frac{E_F[F_m(Y_1) - F(Y_1)]^k}{h_n^k} \right] \\
 &= \frac{1}{m^k h_n^k} \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} m^i E_G \left[\phi_1 \left(\frac{t-1+F(Y_1)}{h_n} \right) \omega_i(F(Y_1)) \right] \\
 &= \frac{1}{m^k h_n^k} \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} m^i \int_{-\infty}^{\infty} \phi_1 \left(\frac{t-1+F(y)}{h_n} \right) \omega_i(F(y)) g(y) dy \\
 &= \frac{1}{m^k h_n^{k-1}} \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} m^i \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_1(x) \tilde{\omega}_i(t, x h_n) r(t - x h_n) dx,
 \end{aligned} \tag{57}$$

where

$$r(t) = R'(t) = \frac{g(F^{-1}(1-t))}{f(F^{-1}(1-t))} \tag{58}$$

and $\tilde{\omega}_i(t, x h_n) = \omega_i(1-t+x h_n) = \sum_{j=0}^k a_{i,j}(t) (x h_n)^j$ for some $a_{i,j}(t) \in \mathbb{R}$. Expanding $r(t - x h_n)$ in a Taylor series at t , we get

$$\begin{aligned}
 &\int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_1(x) \tilde{\omega}_i(t, x h_n) r(t - x h_n) dx \\
 &= r(t) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_1(x) \tilde{\omega}_i(t, x h_n) dx - r'(t) h_n \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x \phi_1(x) \tilde{\omega}_i(t, x h_n) dx + o(h_n) \\
 &= r(t) \sum_{j=0}^k a_{i,j}(t) h_n^j \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x^j \phi_1(x) dx + O(h_n) \\
 &= r(t) a_{i,0}(t) \int_{-1}^1 \phi_1(x) dx + r(t) \sum_{j=1}^k a_{i,j}(t) h_n^j \int_{-1}^1 x^j \phi_1(x) dx + O(h_n).
 \end{aligned}$$

The above integral is finite for every $i = 1, 2, \dots, \lfloor \frac{k}{2} \rfloor$ and, if ϕ_1 is odd and in consequence $\int_{-1}^1 \phi_1(x) dx = 0$, it is of order $O(h_n)$. This, in combination with (57), proves statement (1) of the lemma. To prove statement (2) let us denote $DF(x, y) = F(y) - F(x)$ and $DF_m(x, y) = F_m(y) - F_m(x)$, $x, y \in \mathbb{R}$ and assume that $x < y$.

Then

$$\begin{aligned}
 & \mathbb{E}_F[F_m(x) - F(x)]^k [F_m(y) - F(y)]^l \\
 &= \frac{1}{m^{k+l}} \mathbb{E}_F[mF_m(x) - mF(x)]^k [mF_m(y) - mF(y)]^l \\
 &= \frac{1}{m^{k+l}} \mathbb{E}_F[mF_m(x) - mF(x)]^k \\
 & \quad [mDF_m(x, y) - mDF(x, y) + mF_m(x) - mF(x)]^l \\
 &= \frac{1}{m^{k+l}} \sum_{s=0}^l \binom{l}{s} \mathbb{E}_F[mF_m(x) - mF(x)]^{k+l-s} [mDF_m(x, y) - mDF(x, y)]^s.
 \end{aligned}$$

The random vector $(mF_m(x), mDF_m(x, y), m - mF_m(x) - mDF_m(x, y))$ has a multinomial distribution $\mathcal{M}_3(m, \mathbf{p})$, where $\mathbf{p} = (F(x), DF(x, y), 1 - F(x) - DF(x, y))$. Therefore, from Lemma 2 we conclude that all expectations in the sum on the right side of the above equation, are polynomials of degree $\lfloor \frac{k+l}{2} \rfloor$ in m . Hence

$$\mathbb{E}_F[F_m(x) - F(x)]^k [F_m(y) - F(y)]^l = O\left(\frac{m^{\lfloor \frac{k+l}{2} \rfloor}}{m^{k+l}}\right). \quad (59)$$

A similar arguments leads to the same conclusion when $y < x$, so equality (59) is true for any $x, y \in \mathbb{R}$. Therefore

$$\begin{aligned}
 & \left| \mathbb{E}_{F,G} \left[\phi_1(T_1) \phi_2(T_2) (T_{1,m} - T_1)^k (T_{2,m} - T_2)^l \right] \right| \\
 &= \left| \mathbb{E}_G \left[\prod_{i=1}^2 \phi_i \left(\frac{t-1+F(Y_i)}{h_n} \right) \frac{\mathbb{E}_F[F_m(Y_1) - F(Y_1)]^k [F_m(Y_2) - F(Y_2)]^l}{h_n^{k+l}} \right] \right| \\
 &\leq \frac{Cm^{\lfloor \frac{k+l}{2} \rfloor}}{m^{k+l} h_n^{k+l}} \mathbb{E}_G \left[\prod_{i=1}^2 \left| \phi_i \left(\frac{t-1+F(Y_i)}{h_n} \right) \right| \right],
 \end{aligned} \quad (60)$$

where $C > 0$ is some constant. With respect to the last expectation, we have

$$\begin{aligned}
 & \mathbb{E}_G \left[\prod_{i=1}^2 \left| \phi_i \left(\frac{t-1+F(Y_i)}{h_n} \right) \right| \right] = \prod_{i=1}^2 \mathbb{E}_G \left| \phi_i \left(\frac{t-1+F(Y_1)}{h_n} \right) \right| \\
 &= \prod_{i=1}^2 \int_{-\infty}^{\infty} \left| \phi_i \left(\frac{t-1+F(y)}{h_n} \right) \right| g(y) dy = h_n^2 \prod_{i=1}^2 \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} |\phi_i(x)| r(t - xh_n) dx \\
 &\leq h_n^2 \prod_{i=1}^2 \int_{-1}^1 |\phi_i(x)| r(t - xh_n) dx \leq h_n^2 \phi_1^* \phi_2^* \left(\int_{-1}^1 r(t - xh_n) dx \right)^2 \\
 &= h_n^2 \phi_1^* \phi_2^* (2r(t) + o(1))^2 = 4h_n^2 \phi_1^* \phi_2^* r^2(t) + o(h_n^2),
 \end{aligned}$$

where $r(t)$ is given by (58) and $\phi_j^* = \sup_{x \in [-1, 1]} |\phi_j(x)| < \infty$, $j = 1, 2$. This, in combination with (60), completes the proof of statement (2) of the lemma. \square

Lemma 4 Let Y_1 and Y_2 be independent random variables from a continuous distribution function G and let F be a continuous distribution function. Let ϕ_1 and ϕ_2 be continuous functions supported on $[-1, 1]$ and let $R(s) = 1 - G(F^{-1}(1 - s))$. Assume that $R''(s)$ exists for s near $t \in (0, 1)$ and $R''(s)$ is continuous at $s = t$. Denote $T_i = \frac{t-1+F(Y_i)}{h_n}$, $i = 1, 2$ and $H(x, y, z) = \sum_{k=1}^{k_0} c_k x^{\alpha_k} y^{\beta_k} z^{\gamma_k}$, $k_0 \in \mathbb{N}$, $c_k \in \mathbb{R}$, $\alpha_k, \beta_k, \gamma_k \in \mathbb{N}$, $k = 1, 2, \dots, k_0$. Then for any $h_n > 0$ such that $h_n \rightarrow 0$ and for any $k, n \in \mathbb{N}$, we have:

$$\begin{aligned} & E_G [\phi_1(T_1) \phi_2(T_2) H(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2))] \\ &= h_n^2 \eta_{0,1} \eta_{0,2} r^2(t) \sum_{k=1}^{k_0} c_k (1-t)^{\sigma_k} - h_n^3 (\eta_{0,1} \eta_{1,2} + \eta_{0,2} \eta_{1,1}) r(t) r'(t) \sum_{k=1}^{k_0} c_k (1-t)^{\sigma_k} \\ &+ h_n^3 (\eta_2 + \eta_3) r^2(t) \sum_{k=1}^{k_0} c_k \alpha_k (1-t)^{\sigma_k-1} + h_n^3 \eta_{0,2} \eta_{1,1} r^2(t) \sum_{k=1}^{k_0} c_k \beta_k (1-t)^{\sigma_k-1} \\ &+ h_n^3 \eta_{0,1} \eta_{1,2} r^2(t) \sum_{k=1}^{k_0} c_k \gamma_k (1-t)^{\sigma_k-1} + o(h_n^3), \end{aligned}$$

where $r(t) = R'(t) = \frac{g(F^{-1}(1-t))}{f(F^{-1}(1-t))}$, $\eta_{0,i} = \int_{-1}^1 \phi_i(x) dx$, $\eta_{1,i} = \int_{-1}^1 x \phi_i(x) dx$, $i = 1, 2$, $\eta_2 = \int_{-1}^1 \phi_2(u) \int_{-1}^u x \phi_1(x) dx du$, $\eta_3 = \int_{-1}^1 u \phi_2(u) \int_u^1 \phi_1(x) dx du$ and $\sigma_k = \alpha_k + \beta_k + \gamma_k$.

Proof

$$\begin{aligned} I &:= E_G [\phi_1(T_1) \phi_2(T_2) H(F(Y_1 \wedge Y_2), F(Y_1), F(Y_2))] \\ &= \iint_{\mathbb{R}^2} \prod_{l=1}^2 \phi_l \left(\frac{t-1+F(y_l)}{h_n} \right) g(y_l) H(F(y_1 \wedge y_2), F(y_1), F(y_2)) dy_1 dy_2 \\ &= \iint_{A_1} \prod_{l=1}^2 \phi_l \left(\frac{t-1+F(y_l)}{h_n} \right) g(y_l) H(F(y_1), F(y_1), F(y_2)) dy_1 dy_2 \\ &+ \iint_{A_2} \prod_{l=1}^2 \phi_l \left(\frac{t-1+F(y_l)}{h_n} \right) g(y_l) H(F(y_2), F(y_1), F(y_2)) dy_1 dy_2 \\ &:= I_1 + I_2, \end{aligned} \tag{61}$$

where $A_1 = \mathbb{R} \times (-\infty, y_2]$ and $A_2 = \mathbb{R} \times (y_2, \infty)$. Changing variables in the integral I_1 and using the definition of the function H , we get

$$\begin{aligned} I_1 &= \iint_{\widetilde{A}_1} \prod_{l=1}^2 \phi_l(x_l) r(t - x_l h_n) H(1 - t + x_1 h_n, 1 - t + x_1 h_n, 1 - t + x_2 h_n) h_n^2 dx_1 dx_2 \\ &= h_n^2 \sum_{k=1}^{k_0} c_k \iint_{\widetilde{A}_1} \prod_{l=1}^2 \phi_l(x_l) r(t - x_l h_n) (1 - t + x_1 h_n)^{\alpha_k + \beta_k} \\ &\quad (1 - t + x_2 h_n)^{\gamma_k} dx_1 dx_2, \end{aligned}$$

where $\widetilde{A}_1 = \left(\frac{t-1}{h_n}, \frac{t}{h_n}\right) \times \left(\frac{t-1}{h_n}, x_2\right]$ and $r(t)$ is given by (58). Applying Newton's binomial formula to $(1 - t + x_1 h_n)^{\alpha_k + \beta_k}$ and $(1 - t + x_2 h_n)^{\gamma_k}$, we obtain

$$\begin{aligned} I_1 &= h_n^2 \sum_{k=1}^{k_0} c_k \sum_{i=0}^{\gamma_k} \sum_{j=0}^{\alpha_k + \beta_k} \binom{\gamma_k}{i} \binom{\alpha_k + \beta_k}{j} (1 - t)^{\sigma_k - i - j} h_n^{i+j} \\ &\quad \times \iint_{\widetilde{A}_1} x_1^j x_2^i \prod_{l=1}^2 \phi_l(x_l) r(t - x_l h_n) dx_1 dx_2 \\ &= h_n^2 \sum_{k=1}^{k_0} c_k (1 - t)^{\sigma_k} I_{1,0} + h_n^3 \sum_{k=1}^{k_0} c_k (\alpha_k + \beta_k) (1 - t)^{\sigma_k - 1} I_{1,1} \\ &\quad + h_n^3 \sum_{k=1}^{k_0} c_k \gamma_k (1 - t)^{\sigma_k - 1} I_{1,2} + O(h_n^4), \end{aligned} \tag{62}$$

where

$$\begin{aligned} I_{1,0} &= \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_2(x_2) r(t - x_2 h_n) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_1(x_1) r(t - x_1 h_n) dx_1 dx_2, \\ I_{1,1} &= \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_2(x_2) r(t - x_2 h_n) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_1 \phi_1(x_1) r(t - x_1 h_n) dx_1 dx_2, \\ I_{1,2} &= \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_2 \phi_2(x_2) r(t - x_2 h_n) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_1(x_1) r(t - x_1 h_n) dx_1 dx_2. \end{aligned}$$

Analogously, one can check that in the case of the integral I_2 , we have

$$\begin{aligned} I_2 &= h_n^2 \sum_{k=1}^{k_0} c_k (1 - t)^{\sigma_k} I_{2,0} + h_n^3 \sum_{k=1}^{k_0} c_k \beta_k (1 - t)^{\sigma_k - 1} I_{2,1} \\ &\quad + h_n^3 \sum_{k=1}^{k_0} c_k (\alpha_k + \gamma_k) (1 - t)^{\sigma_k - 1} I_{2,2} + O(h_n^4), \end{aligned} \tag{63}$$

where

$$\begin{aligned} I_{2,0} &= \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_2(x_2) r(t - x_2 h_n) \int_{x_2}^{\frac{t}{h_n}} \phi_1(x_1) r(t - x_1 h_n) dx_1 dx_2, \\ I_{2,1} &= \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_2(x_2) r(t - x_2 h_n) \int_{x_2}^{\frac{t}{h_n}} x_1 \phi_1(x_1) r(t - x_1 h_n) dx_1 dx_2, \\ I_{2,2} &= \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_2 \phi_2(x_2) r(t - x_2 h_n) \int_{x_2}^{\frac{t}{h_n}} \phi_1(x_1) r(t - x_1 h_n) dx_1 dx_2. \end{aligned}$$

Combining (61), (62) and (63), after simplification, we have

$$\begin{aligned} I &= I_1 + I_2 = h_n^2 \sum_{k=1}^{k_0} c_k (1-t)^{\sigma_k} (I_{1,0} + I_{2,0}) \\ &\quad + h_n^3 \sum_{k=1}^{k_0} c_k \alpha_k (1-t)^{\sigma_k-1} (I_{1,1} + I_{2,2}) + h_n^3 \sum_{k=1}^{k_0} c_k \beta_k (1-t)^{\sigma_k-1} (I_{1,1} + I_{2,1}) \\ &\quad + h_n^3 \sum_{k=1}^{k_0} c_k \gamma_k (1-t)^{\sigma_k-1} (I_{1,2} + I_{2,2}) + O(h_n^4). \end{aligned} \quad (64)$$

Note that

$$\begin{aligned} I_{1,0} + I_{2,0} &= \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_1(x_1) r(t - x_1 h_n) dx_1 \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_2(x_2) r(t - x_2 h_n) dx_2, \\ I_{1,1} + I_{2,1} &= \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_1 \phi_1(x_1) r(t - x_1 h_n) dx_1 \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_2(x_2) r(t - x_2 h_n) dx_2, \\ I_{1,2} + I_{2,2} &= \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_1(x_1) r(t - x_1 h_n) dx_1 \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_2 \phi_2(x_2) r(t - x_2 h_n) dx_2. \end{aligned} \quad (65)$$

Expanding $r(t - x_i h_n)$, $i = 1, 2$, in a Taylor series at t in the above integrals, we obtain

$$\begin{aligned} &\int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_i(x_i) r(t - x_i h_n) dx_i \\ &= r(t) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_i(x_i) dx_i - r'(t) h_n \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_i \phi_i(x_i) dx_i + o(h_n) \\ &= r(t) \int_{-1}^1 \phi_i(x_i) dx_i - r'(t) h_n \int_{-1}^1 x_i \phi_i(x_i) dx_i + o(h_n) \\ &= r(t) \eta_{0,i} - h_n r'(t) \eta_{1,i} + o(h_n), \quad i = 1, 2, \end{aligned}$$

and

$$\begin{aligned} \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_i \phi_i(x_i) r(t - x_i h_n) dx_i &= r(t) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_i \phi_i(x_i) dx_i + o(1) \\ &= r(t) \int_{-1}^1 x_i \phi_i(x_i) dx_i + o(1) = r(t) \eta_{1,i} + o(1), \quad i = 1, 2, \end{aligned}$$

where $\eta_{0,i}$, $\eta_{1,i}$, $i = 1, 2$, are define in the lemma. In consequence, substituting the obtained expansions into (65), we get

$$\begin{aligned} I_{1,0} + I_{2,0} &= \eta_{0,1} \eta_{0,2} r^2(t) - h_n (\eta_{0,1} \eta_{1,2} + \eta_{0,2} \eta_{1,1}) r(t) r'(t) + o(h_n), \\ I_{1,1} + I_{2,1} &= \eta_{0,2} \eta_{1,1} r^2(t) + o(1), \\ I_{1,2} + I_{2,2} &= \eta_{1,2} \eta_{0,1} r^2(t) + o(1). \end{aligned} \quad (66)$$

Now we need to check that

$$I_{1,1} + I_{2,2} = (\eta_2 + \eta_3) r^2(t) + o(1). \quad (67)$$

Indeed

$$\begin{aligned} I_{1,1} + I_{2,2} &= \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_2(x_2) r(t - x_2 h_n) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_1 \phi_1(x_1) r(t - x_1 h_n) dx_1 dx_2 \\ &\quad + \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_2 \phi_2(x_2) r(t - x_2 h_n) \int_{x_2}^{\frac{t}{h_n}} \phi_1(x_1) r(t - x_1 h_n) dx_1 dx_2 \\ &= r^2(t) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} \phi_2(x_2) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_1 \phi_1(x_1) dx_1 dx_2 \\ &\quad + r^2(t) \int_{\frac{t-1}{h_n}}^{\frac{t}{h_n}} x_2 \phi_2(x_2) \int_{x_2}^{\frac{t}{h_n}} \phi_1(x_1) dx_1 dx_2 + o(1) \\ &= r^2(t) \int_{-1}^1 \phi_2(x_2) \int_{-1}^{x_2} x_1 \phi_1(x_1) dx_1 dx_2 \\ &\quad + r^2(t) \int_{-1}^1 x_2 \phi_2(x_2) \int_{x_2}^1 \phi_1(x_1) dx_1 dx_2 + o(1) \\ &= (\eta_2 + \eta_3) r^2(t) + o(1). \end{aligned}$$

Substituting (66) and (67) into (64), we obtain

$$\begin{aligned}
 I = & h_n^2 \eta_{0,1} \eta_{0,2} r^2(t) \sum_{k=1}^{k_0} c_k (1-t)^{\sigma_k} \\
 & - h_n^3 (\eta_{0,1} \eta_{1,2} + \eta_{0,2} \eta_{1,1}) r(t) r'(t) \sum_{k=1}^{k_0} c_k (1-t)^{\sigma_k} \\
 & + h_n^3 (\eta_2 + \eta_3) r^2(t) \sum_{k=1}^{k_0} c_k \alpha_k (1-t)^{\sigma_k-1} + h_n^3 \eta_{0,2} \eta_{1,1} r^2(t) \sum_{k=1}^{k_0} c_k \beta_k (1-t)^{\sigma_k-1} \\
 & + h_n^3 \eta_{1,2} \eta_{0,1} r^2(t) \sum_{k=1}^{k_0} c_k \gamma_k (1-t)^{\sigma_k-1} + o(h_n^3),
 \end{aligned}$$

which completes the proof. \square

References

- Altman N, Leger C (1995) Bandwidth selection for kernel distribution function estimation. *J Stat Plan Inference* 46:195–214
- Bowman A, Hall P, Prvan T (1998) Bandwidth selection for the smoothing of distribution functions. *Biometrika* 85:799–808
- Du P, Tang L (2009) Transformation-invariant and nonparametric monotone smooth estimation of ROC curves. *Stat Med* 28:349–359
- Gu J, Ghosal S, Roy A (2008) Bayesian bootstrap estimation of ROC curve. *Stat Med* 27:5407–5420
- Hall PG, Hyndmann RJ (2003) Improved methods for bandwidth selection when estimating ROC curves. *Stat Probab Lett* 64:181–189
- Hsieh F, Turnbull BW (1996) Nonparametric and semiparametric estimation of the receiver operating characteristic curve. *Ann Stat* 24:25–40
- Jokiel-Rokita A, Pulit M (2012) Nonparametric estimation of the ROC curve based on smoothed empirical distribution functions. *Stat Comput* 23:703–712
- Krzanowski WJ, Hand DJ (2009) ROC curves for continuous data. Chapman and Hall/CRC, London
- Lloyd CJ (1998) Using smoothed receiver operating characteristic curves to summarize and compare diagnostic systems. *J Am Stat Assoc* 93:1356–1364
- Lloyd CJ, Zhou Y (1999) Kernel estimators of the ROC curve are better than empirical. *Stat Probab Lett* 44:221–228
- Peng L, Zhou XH (2004) Local linear smoothing of receiver operating characteristic (ROC) curves. *J Stat Plan Inference* 118:129–143
- Pepe MS (2003) The statistical evaluation of medical tests for classification and prediction. Oxford University Press, Oxford
- Polansky AM, Baker ER (2000) Multistage plug-in bandwidth selection for kernel distribution function estimates. *J Stat Comput Simul* 65:63–80
- Sarda P (1993) Smoothing parameter selection for smooth distribution functions. *J Stat Plan Inference* 35:65–75
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293
- Tang L, Du P, Wu C (2010) Compare diagnostic tests using transformation-invariant smoothed ROC curves. *J Stat Plan Inference* 140:3540–3551
- Tonini G, Fratto ME, Imperatori M, Pantano F, Vincenzi B, Santini D (2011) Predictive factors of response to treatment in patients with metastatic renal cell carcinoma: new evidence. *Expert Rev Anticancer Ther* 11(6):921–930

- Zhou XH, Harezlak J (2002) Comparison of bandwidth selection methods for kernel smoothing of ROC curves. *Stat Med* 21:2045–2055
- Znidaric M (2009) Asymptotic expansion for inverse moments of binomial and Poisson distributions. *Open Stat Probab J* 1:710
- Zou KH, Hall WJ, Shapiro DE (1997) Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests. *Stat Med* 16:2143–2156